



# Graphs where every $k$ -subset of vertices is an identifying set

Sylvain Gravier, Svante Janson, Tero Laihonen, Sanna Ranto

## ► To cite this version:

Sylvain Gravier, Svante Janson, Tero Laihonen, Sanna Ranto. Graphs where every  $k$ -subset of vertices is an identifying set. *Discrete Mathematics and Theoretical Computer Science*, 2014, Vol. 16 no. 1 (1), pp.73-88. 10.46298/dmtcs.1253 . hal-00362184

**HAL Id: hal-00362184**

**<https://hal.science/hal-00362184>**

Submitted on 22 Jul 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Graphs where every $k$ -subset of vertices is an identifying set

Sylvain Gravier<sup>1</sup>   Svante Janson<sup>2</sup>   Tero Laihonen<sup>3</sup>   Sanna Ranto<sup>4</sup>

<sup>1</sup>CNRS - Institut Fourier, ERT Maths à Modeler, Saint Martin d'Hères, France

<sup>2</sup>Uppsala University, Department of Mathematics, Sweden

<sup>3</sup>Department of Mathematics and Statistics, University of Turku, Finland

<sup>4</sup>Department of Physics and Astronomy, University of Turku, Finland

received 16<sup>th</sup> Feb. 2010, revised 7<sup>th</sup> Feb. 2014, accepted 10<sup>th</sup> Feb. 2014.

Let  $G = (V, E)$  be an undirected graph without loops and multiple edges. A subset  $C \subseteq V$  is called *identifying* if for every vertex  $x \in V$  the intersection of  $C$  and the closed neighbourhood of  $x$  is nonempty, and these intersections are different for different vertices  $x$ . Let  $k$  be a positive integer. We will consider graphs where *every*  $k$ -subset is identifying. We prove that for every  $k > 1$  the maximal order of such a graph is at most  $2k - 2$ . Constructions attaining the maximal order are given for infinitely many values of  $k$ . The corresponding problem of  $k$ -subsets identifying any at most  $\ell$  vertices is considered as well.

**Keywords:** identifying code, extremal graph, strongly regular graph, Plotkin bound

## 1 Introduction

Karpovsky *et al.* introduced identifying sets in [9] for locating faulty processors in multiprocessor systems. Since then identifying sets have been considered in many different graphs (see numerous references in [14]) and they find their motivations, for example, in sensor networks and environmental monitoring [10]. For recent developments see for instance [1, 2].

Let  $G = (V, E)$  be a simple undirected graph where  $V$  is the set of vertices and  $E$  is the set of edges. The adjacency between vertices  $x$  and  $y$  is denoted by  $x \sim y$ , and an edge between  $x$  and  $y$  is denoted by  $\{x, y\}$  or  $xy$ . Suppose  $x, y \in V$ . The (*graphical*) *distance* between  $x$  and  $y$  is the number of edges in any shortest path between these vertices and it is denoted by  $d(x, y)$ . If there is no such path, then  $d(x, y) = \infty$ . We denote by  $N(x)$  the set of vertices adjacent to  $x$  (*neighbourhood*) and the *closed neighbourhood* of a vertex  $x$  is  $N[x] = \{x\} \cup N(x)$ . The closed neighbourhood within radius  $r$  centered at  $x$  is denoted by  $N_r[x] = \{y \in V \mid d(x, y) \leq r\}$ . We denote further  $S_r(x) = \{y \in V \mid d(x, y) = r\}$ . Moreover, for  $X \subseteq V$ ,  $N_r[X] = \cup_{x \in X} N_r[x]$ . For  $C \subseteq V$ ,  $X \subseteq V$ , and  $x \in V$  we denote

$$I_r(C; x) = I_r(x) = N_r[x] \cap C$$

<sup>(i)</sup> Emails: Sylvain.Gravier@ujf-grenoble.fr, svante.janson@math.uu.se, terolai@utu.fi, samano@utu.fi

and

$$I_r(C; X) = I_r(X) = N_r[X] \cap C = \bigcup_{x \in X} I_r(C; x).$$

If  $r = 1$ , we drop it from the notations. When necessary, we add a subscript  $G$ . We also write, for example,  $N[x, y]$  and  $I(C; x, y)$  for  $N[\{x, y\}]$  and  $I(C; \{x, y\})$ . The *symmetric difference* of two sets is

$$A \triangle B = (A \setminus B) \cup (B \setminus A).$$

The cardinality of a set  $X$  is denoted by  $|X|$ ; we will also write  $|G|$  for the order  $|V|$  of a graph  $G = (V, E)$ . The *degree* of a vertex  $x$  is  $\deg(x) = |N(x)|$ . Moreover,  $\delta_G = \delta = \min_{x \in V} \deg(x)$  and  $\Delta_G = \Delta = \max_{x \in V} \deg(x)$ . The *diameter* of a graph  $G = (V, E)$  is  $\text{diam}(G) = \max\{d(x, y) \mid x, y \in V\}$ .

We say that a vertex  $x \in V$  *dominates* a vertex  $y \in V$  if and only if  $y \in N[x]$ . As well we can say that a vertex  $y$  is *dominated* by  $x$  (or vice versa). A subset  $C$  of vertices  $V$  is called a *dominating set* (or *dominating*) if  $\cup_{c \in C} N[c] = V$ .

**Definition 1** A subset  $C$  of vertices of a graph  $G = (V, E)$  is called  $(r, \leq \ell)$ -identifying (or an  $(r, \leq \ell)$ -identifying set) if for all  $X, Y \subseteq V$  with  $|X| \leq \ell$ ,  $|Y| \leq \ell$ ,  $X \neq Y$  we have

$$I_r(C; X) \neq I_r(C; Y).$$

If  $r = 1$  and  $\ell = 1$ , then we speak about an identifying set.

The idea behind identification is that we can uniquely determine the subset  $X$  of vertices of a graph  $G = (V, E)$  by knowing only  $I_r(C; X)$  — provided that  $|X| \leq \ell$  and  $C \subseteq V$  is an  $(r, \leq \ell)$ -identifying set.

**Definition 2** Let, for  $n \geq k \geq 1$  and  $\ell \geq 1$ ,  $\mathfrak{Gr}(n, k, \ell)$  be the set of graphs on  $n$  vertices such that every  $k$ -element set of vertices is  $(1, \leq \ell)$ -identifying. Moreover, we denote  $\mathfrak{Gr}(n, k, 1) = \mathfrak{Gr}(n, k)$  and  $\mathfrak{Gr}(k) = \bigcup_{n \geq k} \mathfrak{Gr}(n, k)$ .

In other words, in a sensor network which is modeled by a graph in the class  $\mathfrak{Gr}(n, k, \ell)$  we can choose freely  $k$  sensors [10] i.e. vertices to locate any  $\ell$  objects in vertices.

**Example 3** (i) For every  $\ell \geq 1$ , an empty graph  $E_n = (\{1, \dots, n\}, \emptyset)$  belongs to  $\mathfrak{Gr}(n, k, \ell)$  if and only if  $k = n$ .

(ii) A cycle  $C_n$  ( $n \geq 4$ ) belongs to  $\mathfrak{Gr}(n, k)$  if and only if  $n - 1 \leq k \leq n$ . A cycle  $C_n$  with  $n \geq 7$  is in  $\mathfrak{Gr}(n, n, 2)$ .

(iii) A path  $P_n$  of  $n$  vertices ( $n \geq 3$ ) belongs to  $\mathfrak{Gr}(n, k)$  if and only if  $k = n$ .

(iv) A complete bipartite graph  $K_{n,m}$  ( $n+m \geq 4$ ) is in  $\mathfrak{Gr}(n+m, k)$  if and only if  $n+m-1 \leq k \leq n+m$ .

(v) In particular, a star  $S_n = K_{1,n-1}$  ( $n \geq 4$ ) is in  $\mathfrak{Gr}(n, k)$  if and only if  $n - 1 \leq k \leq n$ .

(vi) The complete graph  $K_n$  ( $n \geq 2$ ) is not in  $\mathfrak{Gr}(n, k)$  for any  $k$ .

We are interested in the maximum number  $n$  of vertices which can be reached by a given  $k$ . We study mainly the case  $\ell = 1$  and define

$$\Xi(k) = \max\{n : \mathfrak{Gr}(n, k) \neq \emptyset\}. \quad (1)$$

Conversely, the question is for a given graph on  $n$  vertices what is the smallest number  $k$  such that every  $k$ -subset of vertices is an identifying set (or a  $(1, \leq \ell)$ -identifying set). (Note that even if we take  $k = n$ , there are graphs on  $n$  vertices that do not belong to  $\mathfrak{Gr}(n, n)$ , for example the complete graph  $K_n$ ,  $n \geq 2$ .) The ratio  $n/k$  is called the *rate*.

In particular, we are interested in the asymptotics as  $k \rightarrow \infty$ . Combining Theorem 17 and Corollary 26, we obtain the following, which in particular shows that the rate is always less than 2.

**Theorem 4**  $\Xi(k) \leq 2k - 2$  for all  $k \geq 2$ , and  $\lim_{k \rightarrow \infty} \frac{\Xi(k)}{k} = 2$ .

We will see in Section 4 that  $\Xi(k) = 2k - 2$  for infinitely many  $k$ .

We give some basic results in Section 2 and study small  $k$  in Section 2.1 where we give a complete description of the sets  $\mathfrak{Gr}(k)$  for  $k \leq 4$ . In Section 3 we give an upper bound, which bases on a relation with error-correcting codes. We consider strongly regular graphs and some modifications of them in Section 4; this provides us with examples (e.g., Paley graphs) that attain or almost attain the upper bound in Theorem 4. In Section 5 we give results for the case  $\ell \geq 2$ .

## 2 Basic results

We begin with some simple consequences of the definition. We omit the simple proofs.

**Lemma 5** If  $G = (V, E) \in \mathfrak{Gr}(n, k, \ell)$ , then every induced subgraph  $G[A]$ , where  $A \subseteq V$ , of order  $|A| = m \geq k$  belongs to  $\mathfrak{Gr}(m, k, \ell)$ .

**Lemma 6** If  $G$  has connected components  $G_i$ ,  $i = 1, \dots, m$ , with  $|G| = n$  and  $|G_i| = n_i$ , then  $G \in \mathfrak{Gr}(n, k, \ell)$  if and only if  $G_i \in \mathfrak{Gr}(n_i, k + n_i - n, \ell)$  for every  $i$ . In other words,  $G_i \in \mathfrak{Gr}(n_i, k_i, \ell)$  with  $n_i - k_i = n - k$ .

A graph  $G$  belongs to  $\mathfrak{Gr}(n, k, \ell)$  if and only if every  $k$ -subset intersects every symmetric difference of the neighbourhoods of two sets that are of size at most  $\ell$ . Equivalently,  $G \in \mathfrak{Gr}(n, k, \ell)$  if and only if the complement of every such symmetric difference of two neighbourhoods contains less than  $k$  vertices. We state this as a theorem.

**Theorem 7** Let  $G = (V, E)$  and  $|V| = n$ . A graph  $G$  belongs to  $\mathfrak{Gr}(n, k, \ell)$  if and only if

$$n - \min_{\substack{X, Y \subseteq V \\ X \neq Y \\ |X|, |Y| \leq \ell}} \{|N[X] \triangle N[Y]|\} \leq k - 1. \quad (2)$$

Now take  $\ell = 1$ , and consider  $\mathfrak{Gr}(n, k)$ . The characterization in Theorem 7 can be written as follows, since  $X$  and  $Y$  either are empty or singletons.

**Corollary 8** Let  $G = (V, E)$  and  $|V| = n$ . A graph  $G$  belongs to  $\mathfrak{Gr}(n, k)$  if and only if

- (i)  $\delta_G \geq n - k$ , and

$$(ii) \max_{x,y \in V, x \neq y} \{|N[x] \cap N[y]| + |V \setminus (N[x] \cup N[y])|\} \leq k - 1.$$

In particular, if  $G \in \mathfrak{Gr}(n, k)$  then every vertex is dominated by every choice of a  $k$ -subset, and for all distinct  $x, y \in V$  we have  $|N[x] \cap N[y]| \leq k - 1$ .

**Example 9** Let  $G$  be the 3-dimensional cube, with 8 vertices. Then  $|N[x]| = 4$  for every vertex  $x$ , and  $|N[x] \triangle N[y]|$  is 4 when  $d(x, y) = 1$ , 4 when  $d(x, y) = 2$ , and 8 when  $d(x, y) = 3$ . Hence, Theorem 7 shows that  $G \in \mathfrak{Gr}(8, 5)$ .

**Lemma 10** Let  $G_0 = (V_0, E_0) \in \mathfrak{Gr}(n_0, k_0)$  and let  $G = (V_0 \cup \{a\}, E_0 \cup \{\{a, x\} \mid x \in V_0\})$  for a new vertex  $a \notin V_0$ . In words, we add a vertex and connect it to all other vertices. Then  $G \in \mathfrak{Gr}(n_0 + 1, k_0 + 1)$  if (and only if)  $|N_{G_0}[x]| \leq k_0 - 1$  for every  $x \in V_0$ , or, equivalently,  $\Delta_{G_0} \leq k_0 - 2$ .

**Proof:** An immediate consequence of Theorem 7 (or Corollary 8).  $\square$

**Example 11** If  $G_0$  is the 3-dimensional cube in Example 9, which belongs to  $\mathfrak{Gr}(8, 5)$  and is regular with degree  $3 = 5 - 2$ , then Lemma 10 yields a graph  $G \in \mathfrak{Gr}(9, 6)$ .  $G$  can be regarded as a cube with centre.

## 2.1 Small $k$

**Example 12** For  $k = 1$ , it is easily seen that  $\mathfrak{Gr}(n, 1) = \emptyset$  for  $n \geq 2$ , and thus  $\mathfrak{Gr}(1) = \{K_1\}$  and  $\Xi(1) = 1$ .

**Example 13** Let  $k = 2$ . If  $G \in \mathfrak{Gr}(2)$ , then  $G$  cannot contain any edge  $xy$ , since then  $N[x] \cap \{x, y\} = \{x, y\} = N[y] \cap \{x, y\}$ , so  $\{x, y\}$  does not separate  $\{x\}$  and  $\{y\}$ . Consequently,  $G$  has to be an empty graph  $E_n$ , and then  $\delta_G = 0$  and Corollary 8(i) (or Example 3(i)) shows that  $n = k = 2$ . Thus  $\mathfrak{Gr}(2) = \{E_2\}$  and  $\Xi(2) = 2$ .

**Example 14** Let  $k = 3$ . First, assume  $n = |G| = 3$ . There are only four graphs  $G$  with  $|G| = 3$ , and it is easily checked that  $E_3, P_3 \in \mathfrak{Gr}(3, 3)$  (Example 3(i)(iii)), while  $C_3 = K_3 \notin \mathfrak{Gr}(3, 3)$  (Example 3(vi)) and a disjoint union  $K_1 \cup K_2 \notin \mathfrak{Gr}(3, 3)$ , for example by Lemma 6 since  $K_2 \notin \mathfrak{Gr}(2, 2)$ . Hence  $\mathfrak{Gr}(3, 3) = \{E_3, P_3\}$ .

Next, assume  $n \geq 4$ . Since there are no graphs in  $\mathfrak{Gr}(n_1, k_1)$  if  $n_1 > k_1$  and  $k_1 \leq 2$ , it follows from Lemma 6 that there are no disconnected graphs in  $\mathfrak{Gr}(n, 3)$  for  $n \geq 4$ . Furthermore, if  $G \in \mathfrak{Gr}(n, 3)$ , then every induced subgraph with 3 vertices is in  $\mathfrak{Gr}(3, 3)$  and is thus  $E_3$  or  $P_3$ ; in particular,  $G$  contains no triangle.

If  $G \in \mathfrak{Gr}(4, 3)$ , it follows easily that  $G$  must be  $C_4$  or  $S_4$ , and indeed these belong to  $\mathfrak{Gr}(4, 3)$  by Example 3(ii)(v). Hence  $\mathfrak{Gr}(4, 3) = \{C_4, S_4\}$ .

Next, assume  $G \in \mathfrak{Gr}(5, 3)$ . Then every induced subgraph with 4 vertices is in  $\mathfrak{Gr}(4, 3)$  and is thus  $C_4$  or  $S_4$ . Moreover, by Corollary 8,  $\delta_G \geq 5 - 3 = 2$ . However, if we add a vertex to  $C_4$  or  $S_4$  such that the degree condition  $\delta_G \geq 2$  is satisfied and we do not create a triangle we get  $K_{2,3}$  – a complete bipartite graph, and we know already  $K_{2,3} \notin \mathfrak{Gr}(5, 3)$  (Example 3(iv)). Consequently  $\mathfrak{Gr}(5, 3) = \emptyset$ , and thus  $\mathfrak{Gr}(n, 3) = \emptyset$  for all  $n \geq 5$ .

Consequently,  $\mathfrak{Gr}(3) = \mathfrak{Gr}(3, 3) \cup \mathfrak{Gr}(4, 3) = \{E_3, P_3, S_4, C_4\}$  and  $\Xi(3) = 4$ .

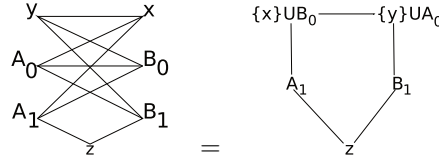
**Example 15** Let  $k = 4$ . First, it follows easily from Lemma 6 and the descriptions of  $\mathfrak{Gr}(j)$  for  $j \leq 3$  above that the only disconnected graphs in  $\mathfrak{Gr}(4)$  are  $E_4$  and the disjoint union  $P_3 \cup K_1$ ; in particular, every graph in  $\mathfrak{Gr}(n, 4)$  with  $n \geq 5$  is connected.

Next, if  $G \in \mathfrak{Gr}(n, 4)$ , there cannot be a triangle in  $G$  because otherwise if a 4-subset includes the vertices of a triangle, one more vertex cannot separate the vertices of the triangle from each other. (Cf. Lemma 19.)

For  $n = 4$ , the only connected graphs of order 4 that do not contain a triangle are  $C_4$ ,  $P_4$  and  $S_4$ , and these belong to  $\mathfrak{Gr}(4, 4)$  by Example 3(ii)(iii)(v). Hence  $\mathfrak{Gr}(4, 4) = \{C_4, P_4, S_4, E_4, P_3 \cup K_1\}$ .

Now assume that  $G \in \mathfrak{Gr}(n, 4)$  with  $n \geq 5$ .

(i) Suppose first that a graph  $K_1 \cup K_2 = (\{x, y, z\}, \{\{x, y\}\})$  is an induced subgraph of  $G$ . Then all the other vertices of  $G$  are adjacent to either  $x$  or  $y$  but not both, since otherwise there would be an induced triangle or an induced  $E_2 \cup K_2$  or  $K_2 \cup K_2$ , and these do not belong to  $\mathfrak{Gr}(4, 4)$ . Let  $A = N(x) \setminus \{y\}$  and  $B = N(y) \setminus \{x\}$ , so we have a partition of the vertex set as  $\{x, y, z\} \cup A \cup B$ . There can be further edges between  $A$  and  $B$ ,  $z$  and  $A$ ,  $z$  and  $B$  but not inside  $A$  and  $B$ . Let  $A = A_0 \cup A_1$  and  $B = B_0 \cup B_1$ , where  $A_1 = \{a \in A \mid a \sim z\}$ ,  $A_0 = A \setminus A_1$  and  $B_1 = \{b \in B \mid b \sim z\}$ ,  $B_0 = B \setminus B_1$ . If  $a \in A_0$  and  $b \in B$ , then the 4-subset  $\{a, b, x, z\}$  does not distinguish  $a$  and  $x$  unless  $a \sim b$ . Similarly, if  $a \in A$  and  $b \in B_0$ , then  $a \sim b$ . On the other hand, if  $a \in A_1$  and  $b \in B_1$ , then  $a \not\sim b$ , since otherwise  $abz$  would be a triangle. Thus, we have, where one or more of the sets  $A_0, A_1, B_0, B_1$  might be empty, where an edge



is a complete bipartite graph on sets incident to it, and there are no edges inside these sets.

If  $n \geq 6$ , then there are at least two elements in one of the sets  $\{x\} \cup B_0$ ,  $\{y\} \cup A_0$ ,  $A_1$  or  $B_1$ . However, these two vertices have the same neighbourhood and hence they cannot be separated by the other  $n - 2 \geq 4$  vertices. Thus,  $n = 5$ .

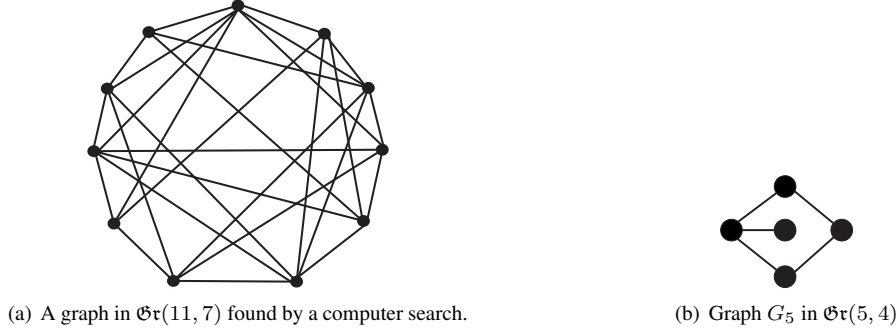
If  $n = 5$ , and both  $A_1$  and  $B_1$  are non-empty, we must have  $A_0 = B_0 = \emptyset$  and  $G = C_5$ , which is in  $\mathfrak{Gr}(5, 4)$  by Example 3(ii).

Finally, assume  $n = 5$  and  $A_1 = \emptyset$  (the case  $B_1 = \emptyset$  is the same after relabelling). Then  $B_1$  is non-empty, since  $G$  is connected. If  $B_0$  is non-empty, let  $b_0 \in B_0$  and  $b_1 \in B_1$ , and observe that  $\{x, b_0, b_1, z\}$  does not separate  $z$  and  $b_1$ . Hence  $B_0 = \emptyset$ . We thus have either  $|A_0| = 1$  and  $|B_1| = 1$ , or  $|A_0| = 0$  and  $|B_1| = 2$ , and both cases yield the graph in Figure 1(b) which easily is seen to be in  $\mathfrak{Gr}(5, 4)$ .

(ii) Suppose that there is no induced subgraph  $K_1 \cup K_2$ . Since  $G$  is connected, we can find an edge  $x \sim y$ . Let, as above,  $A = N(x) \setminus \{y\}$  and  $B = N(y) \setminus \{x\}$ . If  $a \in A$  and  $b \in B$  and  $a \not\sim b$ , then  $(\{a, x, b\}, \{\{a, x\}\})$  is an induced subgraph and we are back in case (i). Hence, all edges between sets  $A$  and  $B$  exist and thus, recalling that  $G$  has no triangles,  $G$  is the complete bipartite graph with bipartition  $(A \cup \{y\}, B \cup \{x\})$ . By Example 3(iv), then  $n \leq 5$ . If  $n = 5$ , we get  $G = K_{2,3}$  or  $G = K_{1,4} = S_5$ , which both belong to  $\mathfrak{Gr}(5, 4)$  by Example 3(iv).

We summarize the result in a theorem.

**Theorem 16** We have  $\Xi(4) = 5$ . More precisely,  $\mathfrak{Gr}(4) = \mathfrak{Gr}(4, 4) \cup \mathfrak{Gr}(5, 4)$ , where  $\mathfrak{Gr}(4, 4) = \{C_4, P_4, S_4, E_4, P_3 \cup K_1\}$  and  $\mathfrak{Gr}(5, 4) = \{S_5, C_5, K_{2,3}, G_5\}$  where  $G_5$  is the graph in Figure 1(b).



**Fig. 1:** Examples in  $\mathfrak{G}_r(11, 7)$  and  $\mathfrak{G}_r(5, 4)$ .

Upper and lower bounds for  $\Xi(k)$  for  $1 \leq k \leq 20$  are given in Table 1. Note that we have determined  $\Xi(k)$  exactly for  $k \leq 6$  and for 9, 19, but not for other values of  $k$  when  $k \leq 20$ .

### 3 Upper estimates on the order

In the next theorem we give an upper bound on  $\Xi(k)$ , which is obtained using knowledge on error-correcting codes.

**Theorem 17** *If  $k \geq 2$ , then  $\Xi(k) \leq 2k - 2$ .*

**Proof:** We begin by giving a construction from a graph in  $\mathfrak{G}_r(n, k)$  to error-correcting codes. A non-existence result of error-correcting codes then yields the non-existence of  $\mathfrak{G}_r(n, k)$  graphs of certain parameters. Let  $G = (V, E) \in \mathfrak{G}_r(n, k)$ , where  $V = \{x_1, x_2, \dots, x_n\}$ . We construct  $n + 1$  binary strings  $y_i = (y_{i1}, \dots, y_{in})$  of length  $n$ , for  $i = 0, \dots, n$ , from the sets  $\emptyset = N[\emptyset]$  and  $N[x_i]$  for  $i = 1, \dots, n$  by defining  $y_{0j} = 0$  for all  $j$  and

$$y_{ij} = \begin{cases} 0 & \text{if } x_j \notin N[x_i] \\ 1 & \text{if } x_j \in N[x_i] \end{cases}, \quad 1 \leq i \leq n.$$

Let  $C$  denote the code which consists of these binary strings as codewords. Because  $G \in \mathfrak{G}_r(n, k)$ , the symmetric difference of two closed neighbourhoods  $N[x_i]$  and  $N[x_j]$ , or of one neighbourhood  $N[x_i]$  and  $\emptyset$ , is at least  $n - k + 1$  by (2); in other words, the minimum Hamming distance  $d(C)$  of the code  $C$  is at least  $n - k + 1$ .

We first give a simple proof that  $\Xi(k) \leq 2k - 1$ . Thus, suppose that there is a  $G \in \mathfrak{G}_r(n, k)$  such that  $n = 2k$ . In the corresponding error-correcting code  $C$ , the minimum distance is at least  $d = n - k + 1 = k + 1 > n/2$ . Let the maximum cardinality of the error-correcting codes of length  $n$  and minimum distance at least  $d$  be denoted by  $A(n, d)$ . We can apply the Plotkin bound (see for example [15, Chapter 2, §2]), which says  $A(n, d) \leq 2 \lfloor d/(2d - n) \rfloor$ , when  $2d > n$ . Thus, we have

$$A(n, d) \leq 2 \left\lfloor \frac{k+1}{2} \right\rfloor \leq k+1.$$

Because  $k + 1 < 2k = n < |C|$ , this contradicts the existence of  $C$ . Hence, there cannot exist a graph  $G \in \mathfrak{Gr}(2k, k)$ , and thus  $\mathfrak{Gr}(n, k) = \emptyset$  when  $n \geq 2k$ .

The Plotkin bound is not strong enough to imply  $\Xi(k) \leq 2k - 2$  in general, but we obtain this from the proof of the Plotkin bound as follows. (In fact, for odd  $k$ ,  $\Xi(k) \leq 2k - 2$  follows from the Plotkin bound for an odd minimum distance. We leave this to the reader since the argument below is more general.)

Suppose that  $G = (V, E) \in \mathfrak{Gr}(n, k)$  with  $n = 2k - 1$ . We thus have a corresponding error-correcting code  $C$  with  $|C| = n + 1 = 2k$  and minimum Hamming distance at least  $n - k + 1 = k$ . Hence, letting  $d$  denote the Hamming distance,

$$\sum_{0 \leq i < j \leq n} d(y_i, y_j) \geq \binom{n+1}{2} k = \frac{2k(2k-1)}{2} k = (2k-1)k^2. \quad (3)$$

On the other hand, if there are  $s_m$  strings  $y_i$  with  $y_{im} = 1$ , and thus  $|C| - s_m = 2k - s_m$  strings with  $y_{im} = 0$ , then the number of ordered pairs  $(i, j)$  such that  $y_{im} \neq y_{jm}$  is  $2s_m(2k - s_m)$  and this parabola gives  $2s_m(2k - s_m) \leq 2k^2$ . Hence each bit contributes at most  $k^2$  to the sum in (3), and summing over  $m$  we find

$$\sum_{0 \leq i < j \leq n} d(y_i, y_j) \leq nk^2 = (2k-1)k^2. \quad (4)$$

Consequently, we have equality in (3) and (4), and thus  $d(y_i, y_j) = k$  for all pairs  $(i, j)$  with  $i \neq j$ .

In particular,  $|N[x_i]| = d(y_i, y_0) = k$  for  $i = 1, \dots, n$ , and thus every vertex in  $G$  has degree  $k - 1$ , i.e.,  $G$  is  $(k - 1)$ -regular. Hence,  $2|E| = n(k - 1) = (2k - 1)(k - 1)$ , and  $k$  must be odd.

Further, if  $i \neq j$ , then  $|N[x_i] \triangle N[x_j]| = d(y_i, y_j) = k$ , and since  $N[x_i] \setminus N[x_j]$  and  $N[x_j] \setminus N[x_i]$  have the same size  $k - |N[x_i] \cap N[x_j]|$ , they have both the size  $k/2$  and  $k$  must be even.

This contradiction shows that  $\mathfrak{Gr}(2k - 1, k) = \emptyset$ , and thus  $\Xi(k) \leq 2k - 2$ .  $\square$

The next theorem (which does not use Theorem 17) will lead to another upper bound in Theorem 20. It can be seen as an improvement for the extreme case  $\mathfrak{Gr}(2k - 2, k)$  of Mantel's [16] theorem on existence of triangles in a graph. Note that this result fails for  $k = 5$  by Example 9.

**Theorem 18** *Suppose  $G \in \mathfrak{Gr}(n, k)$  and  $k \geq 6$ . If  $n \geq 2k - 2$ , then there is a triangle in  $G$ .*

**Proof:** Let  $G = (V, E) \in \mathfrak{Gr}(n, k)$ . Suppose to the contrary that there are no triangles in  $G$ . If there is a vertex  $x \in V$  such that  $\deg(x) \geq k + 1$ , then we select in  $N(x)$  a  $k$ -set  $X$  and a vertex  $y$  outside it; since  $X$  has to dominate  $y$ , it is clear that there exists a triangle  $xyz$ . Hence  $\deg(x) \leq k$  for every  $x$ . On the other hand, we know by Corollary 8(i) that for all  $x \in V$   $\deg(x) \geq n - k \geq k - 2$ .

Let  $x \in V$  be a vertex whose degree is minimum. We denote  $V \setminus N[x] = B$  and we use the fact that  $|B| \leq k - 1$ .

1) Suppose first  $\deg(x) = k$ . Because  $\deg(x)$  is minimum we know that for all  $a \in N(x)$ ,  $\deg(a) = k$ . This is possible if and only if  $|B| = k - 1$  and for all  $a \in N(x)$  we have  $B \cap N(a) = B$ . But then in the  $k$ -subset  $C = \{x\} \cup B$  we have  $I(C; a) = I(C; b)$  for all  $a, b \in N(x)$ . This is impossible.

2) Suppose then  $\deg(x) = k - 1$ . If now  $|B| \leq k - 2$  the graph is impossible as in the first case (choose  $C = N[x]$ ). Hence,  $|B| = k - 1$ . For every  $a \in N(x)$  there are at least  $k - 2$  adjacent vertices in  $B$ , and thus at most 1 non-adjacent. This implies that for all  $a, b \in N(x)$ ,  $a \neq b$ , we have  $|N(a) \cap N(b) \cap B| \geq k - 3 \geq 2$ , when  $k \geq 5$ . Hence, by choosing  $a, b \in N(x)$ ,  $a \neq b$ , we have



the  $k$ -subset  $C = \{x\} \cup (N(x) \setminus \{a, b\}) \cup \{c_1, c_2\}$ , where  $c_1, c_2 \in N(a) \cap N(b) \cap B$ . In this  $k$ -subset  $I(C; a) = I(C; b)$ , which is impossible.

3) Suppose finally  $\deg(x) = k - 2$ . Now  $|B| = k - 1$ , otherwise we cannot have  $n \geq 2k - 2$ . If there is  $b \in B$  such that  $|N(b) \cap N(x)| = k - 2$ , then because  $\deg(b) \leq k$  we have  $|N(b) \cap B| \geq 2$  and  $|B \setminus N[b]| \geq k - 4 \geq 2$ , when  $k \geq 6$ . Hence, there are  $c_1, c_2 \in B \setminus N[b]$ ,  $c_1 \neq c_2$ , and in the  $k$ -subset  $C = N(x) \cup \{c_1, c_2\}$  we have  $I(C; x) = I(C; b)$  which is impossible.

Thus, for all  $b \in B$  we have  $|N(b) \cap N(x)| \leq k - 3$ . On the other hand, each of the  $k - 2$  vertices in  $N(x)$  has at least  $k - 3$  adjacent vertices in  $B$ , so the vertices in  $B$  have on the average at least  $(k - 2)(k - 3)/(k - 1) > k - 4$  adjacent vertices in the set  $N(x)$ . Hence, we can find  $b \in B$  such that  $|N(b) \cap N(x)| = k - 3$ . Because  $\deg(b) \geq k - 2$  we have at least one  $b_0 \in B$  such that  $d(b, b_0) = 1$ . Because there are no triangles, each of the  $k - 3$  neighbours of  $b$  in  $N(x)$  is not adjacent with  $b_0$ , and therefore adjacent to at least  $k - 3$  of the  $k - 2$  vertices in  $B \setminus \{b_0\}$ . Hence, for all  $a_1, a_2 \in N(x) \cap N(b)$ ,  $a_1 \neq a_2$ , we have  $|N(a_1) \cap N(a_2) \cap B| \geq k - 4 \geq 2$  when  $k \geq 6$ . In the  $k$ -subset  $C = \{x, b, c_1, c_2\} \cup (N(x) \setminus \{a_1, a_2\})$ , where  $c_1, c_2 \in N(a_1) \cap N(a_2) \cap B$ , we have  $I(C; a_1) = I(C; a_2)$ , which is impossible.  $\square$

**Lemma 19** *If there is a graph  $G \in \mathfrak{Gr}(n, k)$  that contains a triangle, then  $n \leq 3k - 9$ . (In particular,  $k \geq 5$ .)*

**Proof:** Suppose that  $G = (V, E) \in \mathfrak{Gr}(n, k)$  and that there is a triangle  $\{x, y, z\}$  in  $G$ . Let, for  $v, w \in V$ ,  $J_w(v)$  denote the indicator function given by  $J_w(v) = 1$  if  $v \in N[w]$  and  $J_w(v) = 0$  if  $v \notin N[w]$ . Define the set  $M_{xy} = \{v \in V : J_x(v) = J_y(v)\}$ , and  $M'_{xy} = M_{xy} \setminus \{x, y, z\}$ . Since  $M_{xy}$  does not distinguish  $x$  and  $y$ , we have  $|M_{xy}| \leq k - 1$ . Further,  $\{x, y, z\} \subseteq M_{xy}$ , and thus  $|M'_{xy}| \leq k - 4$ . Define similarly  $M_{xz}, M_{yz}, M'_{xz}, M'_{yz}$ ; the same conclusion holds for these.

Since the indicator functions take only two values,  $M_{xy}, M_{xz}$  and  $M_{yz}$  cover  $V$ , and thus

$$n = |V| = |M'_{xy} \cup M'_{xz} \cup M'_{yz} \cup \{x, y, z\}| \leq 3(k - 4) + 3 = 3k - 9.$$

Since  $n \geq k$ , this entails  $3k - 9 \geq k$  and thus  $k \geq 5$ .  $\square$

The following upper bound is generally weaker than Theorem 17, but it gives the optimal result for  $k = 6$ .

**Theorem 20** *Suppose  $k \geq 6$ . Then  $\Xi(k) \leq 3k - 9$ .*

**Proof:** Suppose that  $G \in \mathfrak{Gr}(n, k)$ . If  $G$  does not contain any triangle, then Theorem 18 yields  $n \leq 2k - 3 \leq 3k - 9$ . If  $G$  does contain a triangle, then Lemma 19 yields  $n \leq 3k - 9$ .  $\square$

## 4 Strongly regular graphs

A graph  $G = (V, E)$  is called *strongly regular* with parameters  $(n, t, \lambda, \mu)$  if  $|V| = n$ ,  $\deg(x) = t$  for all  $x \in V$ , any two adjacent vertices have exactly  $\lambda$  common neighbours, and any two nonadjacent vertices have exactly  $\mu$  common neighbours; we then say that  $G$  is an  $(n, t, \lambda, \mu)$ -SRG. See [3] for more information. By [3, Proposition 1.4.1] we know that if  $G$  is an  $(n, t, \lambda, \mu)$ -SRG, then  $n = t + 1 + t(t - 1 - \lambda)/\mu$ .

We give two examples of strongly regular graphs that will be used below.

**Example 21** The well-known Paley graph  $P(q)$ , where  $q$  is a prime power with  $q \equiv 1 \pmod{4}$ , is a  $(q, (q-1)/2, (q-5)/4, (q-1)/4)$ -SRG, see for example [3]. The vertices of  $P(q)$  are the elements of the finite field  $F_q$ , with an edge  $ij$  if and only if  $i - j$  is a non-zero square in the field; when  $q$  is a prime, this means that the vertices are  $\{1, \dots, q\}$  with edges  $ij$  when  $i - j$  is a quadratic residue mod  $q$ .

**Example 22** Another construction of strongly regular graphs uses a regular symmetric Hadamard matrix with constant diagonal (RSHCD) [6], [4], [5]. In particular, in the case (denoted RSHCD+) of a regular symmetric  $n \times n$  Hadamard matrix  $H = (h_{ij})$  with diagonal entries  $+1$  and constant positive row sums  $2m$  (necessarily even when  $n > 1$ ), then  $n = (2m)^2 = 4m^2$  and the graph  $G$  with vertex set  $\{1, \dots, n\}$  and an edge  $ij$  (for  $i \neq j$ ) if and only if  $h_{ij} = +1$  is a  $(4m^2, 2m^2 + m - 1, m^2 + m - 2, m^2 + m)$ -SRG [4, §8D].

It is not known for which  $m$  such RSHCD+ exist (it has been conjectured that any  $m \geq 1$  is possible) but constructions for many  $m$  are known, see [6], [17, V.3] and [5, IV.24.2]. For example, starting with the  $4 \times 4$  RSHCD+

$$H_4 = \begin{pmatrix} 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 \\ -1 & 1 & 1 & 1 \end{pmatrix}$$

its tensor power  $H_4^{\otimes r}$  is an RSHCD+ with  $n = 4^r$ , and thus  $m = 2^{r-1}$ , for any  $r \geq 1$ . This yields a  $(2^{2r}, 2^{2r-1} + 2^{r-1} - 1, 2^{2r-2} + 2^{r-1} - 2, 2^{2r-2} + 2^{r-1})$ -SRG with vertex set  $\{1, 2, 3, 4\}^r$ , where two different vertices  $(i_1, \dots, i_r)$  and  $(j_1, \dots, j_r)$  are adjacent if and only if the number of coordinates  $\nu$  such that  $i_\nu + j_\nu = 5$  is even.

**Theorem 23** A strongly regular graph  $G = (V, E)$  with parameters  $(n, t, \lambda, \mu)$  belongs to  $\mathfrak{S}\mathfrak{r}(n, k)$  if and only if

$$k \geq \max\{n - t, n - 2t + 2\lambda + 3, n - 2t + 2\mu - 1\},$$

or, equivalently,  $t \geq n - k$  and  $2 \max\{\lambda + 1, \mu - 1\} \leq k + 2t - n - 1$ .

**Proof:** An immediate consequence of Theorem 7, since  $|N[x]| = t + 1$  for every vertex  $x$  and  $|N[x] \triangle N[y]|$  equals  $2(t - \lambda - 1)$  when  $x \sim y$  and  $2(t + 1 - \mu)$  when  $x \not\sim y, x \neq y$ .  $\square$

We can extend this construction to other values of  $n$  by modifying the strongly regular graph.

**Theorem 24** If there exists a strongly regular graph with parameters  $(n_0, t, \lambda, \mu)$ , then for every  $i = 0, \dots, n_0 + 1$  there exists a graph in  $\mathfrak{S}\mathfrak{r}(n_0 + i, k_0 + i)$ , where

$$k_0 = \max\{n_0 - t, t, n_0 - 2t + 2\lambda + 3, n_0 - 2t + 2\mu - 1, 2t - 2\lambda - 1, 2t - 2\mu + 2\},$$

provided  $k_0 \leq n_0$ .

**Proof:** For  $i = 0$ , this is a weaker form of Theorem 23. For  $i \geq 1$ , we suppose that  $G_0 = (V_0, E_0)$  is  $(n_0, t, \lambda, \mu)$ -SRG and build a graph  $G_i$  in  $\mathfrak{S}\mathfrak{r}(n_0 + i, k_0 + i)$  from  $G_0$  by adding suitable new vertices and edges.

If  $1 \leq i \leq n_0$ , choose  $i$  different vertices  $x_1, x_2, \dots, x_i$  in  $V_0$ . Construct a new graph  $G_i = (V_i, E_i)$  by taking  $G_0$  and adding to it new vertices  $x'_1, x'_2, \dots, x'_i$  and new edges  $x'_j y$  for  $j \leq i$  and all  $y \notin N_{G_0}(x_j)$ .

First,  $\deg_{G_i}(x) \geq \deg_{G_0}(x) = t$  for  $x \in V_0$  and  $\deg_{G_i}(x') = n_0 - t$  for  $x' \in V'_i = V_i \setminus V_0$ . We proceed to investigate  $N[x] \triangle N[y]$ , and separate several cases.

(i) If  $x, y \in V_0$ , with  $x \neq y$ , then

$$|N[x] \triangle N[y]| \geq |(N[x] \triangle N[y]) \cap V_0| = |N_{G_0}[x] \triangle N_{G_0}[y]|,$$

which equals  $2(t - \lambda - 1)$  if  $x \sim y$  and  $2(t - \mu + 1)$  if  $x \not\sim y$ .

(ii) If  $x \in V_0, y' \in V'_i$ , then, since  $\triangle$  is associative and commutative,

$$|(N[x] \triangle N[y']) \cap V_0| = |(N_{G_0}[x] \triangle (V_0 \triangle N_{G_0}(y)))| = n_0 - |N_{G_0}[x] \triangle N_{G_0}(y)|,$$

which equals  $n_0 - 1$  if  $x = y$ ,  $n_0 - (2t - 2\lambda - 1)$  if  $x \sim y$ , and  $n_0 - (2t - 2\mu + 1)$  if  $x \not\sim y$  and  $x \neq y$ . If  $x \sim y$ , further,  $|(N[x] \triangle N[y']) \cap V'_i| \geq 1$ , since  $y' \notin N[x]$ .

(iii) If  $x', y' \in V'_i$ , with  $x' \neq y'$ , then

$$|(N[x'] \triangle N[y']) \cap V_0| = |(V_0 \setminus N_{G_0}(x)) \triangle (V_0 \setminus N_{G_0}(y))| = |N_{G_0}(x) \triangle N_{G_0}(y)|,$$

which equals  $2(t - \lambda)$  if  $x \sim y$  and  $2(t - \mu)$  if  $x \not\sim y$ . Further,  $|(N[x'] \triangle N[y']) \cap V'_i| = |\{x', y'\}| = 2$ .

Collecting these estimates, we see that  $G_i \in \mathfrak{Gr}(n_0 + i, k_0 + i)$  by Theorem 7 (or Corollary 8) with our choice of  $k_0$ . Note that  $2k_0 \geq (n_0 - 2t + 2\lambda + 3) + (2t - 2\lambda - 1) = n_0 + 2 \geq 3$ , so  $k_0 \geq 2$ .

Finally, for  $i = n_0 + 1$ , we construct  $G_{n_0+1}$  by adding a new vertex to  $G_{n_0}$  and connecting it to all other vertices. The graph  $G_{n_0}$  has by construction maximum degree  $\Delta_{G_{n_0}} = n_0 \leq k_0 + n_0 - 2$ . Hence, Lemma 10 shows that  $G_{n_0+1} \in \mathfrak{Gr}(2n_0 + 1, k_0 + n_0 + 1)$ .  $\square$

We specialize to the Paley graphs, and obtain from Example 21 and Theorems 23–24 the following.

**Theorem 25** *Let  $q$  be an odd prime power such that  $q \equiv 1 \pmod{4}$ .*

(i) *The Paley graph  $P(q) \in \mathfrak{Gr}(q, (q+3)/2)$ .*

(ii) *There exists a graph in  $\mathfrak{Gr}(q+i, (q+3)/2+i)$  for all  $i = 0, 1, \dots, q+1$ .*

Note that the rate  $2q/(q+3)$  for the Paley graphs approaches 2 as  $q \rightarrow \infty$ ; in fact, with  $n = q$  and  $k = (q+3)/2$  we have  $n = 2k - 3$ , almost attaining the bound  $2k - 2$  in Theorem 17. (The Paley graphs thus almost attain the bound in Theorem 17, but never attain it exactly.)

**Corollary 26**  $\Xi(k) \geq 2k - o(k)$  as  $k \rightarrow \infty$ .

**Proof:** Let  $q = p^2$  where (for  $k \geq 6$ )  $p$  is the largest prime such that  $p \leq \sqrt{2k - 3}$ . It follows from the prime number theorem that  $p/\sqrt{2k - 3} \rightarrow 1$  as  $k \rightarrow \infty$ , and thus  $q = 2k - o(k)$ . Hence, if  $k$  is large enough, then  $k \leq q \leq 2k - 3$ , and Theorem 25 shows that  $P(q) \in \mathfrak{Gr}(q, (q+3)/2) \subseteq \mathfrak{Gr}(q, k)$ , so  $\Xi(k) \geq q = 2k - o(k)$ . (Alternatively, we may let  $q$  be the largest prime such that  $q \leq 2k - 3$  and  $q \equiv 1 \pmod{4}$  and use the prime number theorem for arithmetic progressions [8, Chapter 17] to see that then  $q = 2k - o(k)$ .)  $\square$

We turn to the strongly regular graphs constructed in Example 22 and find from Theorem 23 that they are in  $\mathfrak{Gr}(4m^2, 2m^2 + 1)$ , thus attaining the bound in Theorem 17. We state that as a theorem.

**Theorem 27** *The strongly regular graph constructed in Example 22 from an  $n \times n$  RSHCD+ belongs to  $\mathfrak{Gr}(n, n/2 + 1)$ .*

**Corollary 28** *There exist infinitely many integers  $k$  such that  $\Xi(k) = 2k - 2$ .*

**Proof:** If  $k = n/2 + 1$  for an even  $n$  such that there exists an  $n \times n$  RSHCD+, then  $\Xi(k) \geq n = 2k - 2$  by Theorem 27. The opposite inequality is given by Theorem 17. By Example 22, this holds at least for  $k = 2^{2r-1} + 1$  for any  $r \geq 1$ .  $\square$

## 5 On $\mathfrak{Gr}(n, k, \ell)$

In this section we consider  $\mathfrak{Gr}(n, k, \ell)$  for  $\ell \geq 2$ . Let us denote

$$\Xi(k, \ell) = \max\{n : \mathfrak{Gr}(n, k, \ell) \neq \emptyset\}.$$

Trivially, the empty graph  $E_k \in \mathfrak{Gr}(k, k, \ell)$  for any  $\ell \geq 1$ ; thus  $\Xi(k, \ell) \geq k$ .

Note that a graph  $G = (V, E)$  with  $|V| = n$  admits a  $(1, \leq \ell)$ -identifying set  $\iff V$  is  $(1, \leq \ell)$ -identifying  $\iff G \in \mathfrak{Gr}(n, n, \ell)$ .

**Theorem 29** *Suppose that  $G = (V, E) \in \mathfrak{Gr}(n, k, \ell)$ , where  $n > k$  and  $\ell \geq 2$ . Then the following conditions hold:*

- (i) *For all  $x \in V$  we have  $\ell + 1 < n - k + \ell + 1 \leq |N[x]| \leq k - \ell$ . In other words,  $\delta_G \geq n - k + \ell$  and  $\Delta_G \leq k - \ell - 1$ .*
- (ii) *For all  $x, y \in V$ ,  $x \neq y$ ,  $|N[x] \cap N[y]| \leq k - 2\ell + 1$ .*
- (iii)  *$n \leq 2k - 2\ell - 1$  and  $k \geq 2\ell + 2$ .*

**Proof:** (i) Suppose first that there is a vertex  $x \in V$  such that  $|N[x]| \leq n - k + \ell$ . By removing  $n - k$  vertices from  $V$ , starting in  $N[x]$ , we find a  $k$ -subset  $C$  with  $I(C; x) = \{c_1, \dots, c_m\}$  for some  $m \leq \ell$ . If  $m = 0$ , then  $I(C; x) = I(C; \emptyset)$ , which is impossible. If  $1 \leq m < \ell$ , we can arrange (by removing  $x$  first) so that  $x \notin C$ , and thus  $x \notin Y = \{c_1, \dots, c_m\}$ . Then  $I(C; \{x\} \cup Y) = I(C; Y)$ , a contradiction. If  $m = \ell \geq 2$ , we can conversely arrange so that  $x \in C$ , and thus  $x \in I(C; x)$ , say  $c_1 = x$ . Then  $I(C; c_2, \dots, c_m) = I(C; c_1, \dots, c_m)$ , another contradiction. Consequently,  $|N[x]| \geq n - k + \ell + 1$ .

Suppose then  $|N[x]| \geq k - \ell + 1$ . If  $|N[x]| \geq k$ , we can choose a  $k$ -subset  $C$  of  $N[x]$ ; then  $I(C; x) = C = I(C; x, y)$  for any  $y$ , which is impossible. If  $k > |N[x]| \geq k - \ell + 1$ , we can choose a  $k$ -subset  $C = N[x] \cup \{c_1, \dots, c_{k-|N[x]|}\}$ . Choose also  $a \in N(c_1)$  (which is possible because  $\deg(c_1) \geq 1$  by (i)). Now  $I(C; x, c_1, \dots, c_{k-|N[x]|}) = C = I(C; x, a, c_2, \dots, c_{k-|N[x]|})$ , which is impossible.

(ii) Suppose to the contrary that there are  $x, y \in V$ ,  $x \neq y$ , such that  $|N[x] \cap N[y]| \geq k - 2\ell + 2$ . Let  $A = N(y) \setminus N[x]$ . Then, according to (i),  $|A| \leq |N[y] \setminus N[x]| = |N[y]| - |N[x] \cap N[y]| \leq k - \ell - (k - 2\ell + 2) = \ell - 2$ . Since  $k > \ell - 2$  by (i), there is a  $k$ -subset  $C \subseteq V \setminus \{y\}$  such that  $A \subset C$ . Then  $I(C; A \cup \{x, y\}) = I(C; A \cup \{x\})$ , a contradiction.

(iii) An immediate consequence of (i), which implies  $n - k + \ell + 1 \leq k - \ell$  and  $\ell + 1 < k - \ell$ .  $\square$

**Theorem 30** *For  $\ell \geq 2$ ,  $\Xi(k, \ell) \leq \max\{\frac{\ell}{\ell-1}(k-2), k\}$ .*

**Proof:** If  $\Xi(k, \ell) = k$ , there is nothing to prove. Assume then that there exists a graph  $G = (V, E) \in \mathfrak{Gr}(n, k, \ell)$ , where  $n > k$ . By Theorem 29(iii),  $\ell < k/2 < n$ . Let us consider any set of vertices  $Z = \{z_1, z_2, \dots, z_\ell\}$  of size  $\ell$ . We will estimate  $|N[Z]|$  as follows. By Theorem 29(i) we know  $|N[z_1]| \geq n - k + \ell + 1$ . Now  $N[z_1, z_2]$  must contain at least  $n - k + 1$  vertices, which *do not* belong to  $N[z_1]$  due to Theorem 7 which says that  $|N[X] \triangle N[Y]| \geq n - k + 1$ , where we take  $X = \{z_1\}$  and  $Y = \{z_1, z_2\}$ . Analogously, each set  $N[z_1, \dots, z_i]$  ( $i = 2, \dots, \ell$ ) must contain at least  $n - k + 1$  vertices which are not in  $N[z_1, \dots, z_{i-1}]$ . Hence, for the set  $Z$  we have  $|N[Z]| \geq n - k + \ell + 1 + (\ell - 1)(n - k + 1) = \ell(n - k + 2)$ . Since trivially  $|N[Z]| \leq n$ , we have  $(\ell - 1)n \leq \ell(k - 2)$ , and the claim follows.  $\square$

**Corollary 31** For  $\ell \geq 2$ , we have  $\frac{\Xi(k, \ell)}{k} \leq 1 + \frac{1}{\ell - 1}$ .

The next results improve the result of Theorem 30 for  $\ell = 2$ .

**Lemma 32** Assume that  $n > k$ . Let  $G = (V, E)$  belong to  $\mathfrak{Gr}(n, k, 2)$ . Then

$$n + \frac{n - k + 2}{n - 1}(n - k + 3) \leq 2k - 3$$

**Proof:** Suppose  $x \in V$ . Let

$$f(n, k) = \frac{n - k + 2}{n - 1}(n - k + 3).$$

Our aim is first to show that there exists a vertex in  $N(x)$  or in  $S_2(x)$  which dominates at least  $f(n, k)$  vertices of  $N[x]$ . Let

$$\lambda_x = \max\{|N[x] \cap N[a]| \mid a \in N(x)\}.$$

If  $\lambda_x \geq f(n, k)$ , we are already done. But if  $\lambda_x < f(n, k)$ , then we show that there is a vertex in  $S_2(x)$  that dominates at least  $f(n, k)$  vertices of  $N[x]$ . Let us estimate the number of edges between the vertices in  $N(x)$  and in  $S_2(x)$  — we denote this number by  $M$ . By Theorem 29(i), every vertex  $y \in N(x)$  yields at least  $|N[y]| - \lambda_x \geq n - k + 3 - \lambda_x$  such edges and there are at least  $n - k + 2$  vertices in  $N(x)$ . Consequently,  $M \geq (n - k + 2)(n - k + 3 - \lambda_x)$ . On the other hand, again by Theorem 29(i),  $|S_2(x)| \leq n - |N[x]| \leq k - 3$ . Hence, there must exist a vertex in  $S_2(x)$  incident with at least  $M/(k - 3)$  edges whose other endpoint is in  $N(x)$ . Now, if  $\lambda_x < f(n, k)$ , then

$$\frac{M}{k - 3} > \frac{(n - k + 2)(n - k + 3 - f(n, k))}{k - 3} = f(n, k).$$

Hence there exists in this case a vertex in  $S_2(x)$  that is incident to at least  $f(n, k)$  such edges, i.e., it dominates at least  $f(n, k)$  vertices in  $N(x)$ .

In any case there thus exists  $z \neq x$  such that  $|N[x] \cap N[z]| \geq f(n, k)$ . Let  $C = (N[x] \cap N[z]) \cup (V \setminus N[x])$ . Then  $I(C; x, z) = I(C; z)$ , so  $C$  is not  $(1, \leq 2)$ -identifying and thus  $|C| < k$ . Hence, using Theorem 29(i),

$$k - 1 \geq |C| \geq f(n, k) + n - |N[x]| \geq f(n, k) + n - (k - 2),$$

and thus  $n + f(n, k) \leq 2k - 3$  as asserted.  $\square$

**Theorem 33** *If  $k \leq 5$ , then  $\Xi(k, 2) = k$ . If  $k \geq 6$ , then*

$$\Xi(k, 2) < \left(1 + \frac{1}{\sqrt{2}}\right)(k - 2) + \frac{1}{4}.$$

**Proof:** Let  $n = \Xi(k, 2)$ , and let  $m = k - 2$ . If  $n > k$ , then  $k \geq 6$  by Theorem 29(iii); hence  $n = k$  when  $k \leq 5$ . Further, still assuming  $n > k$ , Lemma 32 yields

$$n + \frac{(n - m)(n - m + 1)}{n - 1} \leq 2m + 1$$

or

$$0 \geq n(n - 1) + (n - m)^2 + n - m - (2m + 1)(n - 1) = 2\left(n - \left(m + \frac{1}{4}\right)\right)^2 - m^2 + \frac{7}{8}.$$

Hence,  $n - \left(m + \frac{1}{4}\right) < m/\sqrt{2}$ . □

**Corollary 34** *For  $\ell = 2$ , we have  $\Xi(k, 2)/k \leq 1 + \frac{1}{\sqrt{2}}$ .*

**Problem 35** *What is  $\limsup_{k \rightarrow \infty} \Xi(k, \ell)/k$  for  $\ell \geq 2$ ? In particular, is  $\limsup_{k \rightarrow \infty} \Xi(k, \ell)/k > 1$ ?*

The following theorem implies that for any  $\ell \geq 2$  there exist graphs in  $\mathfrak{Gr}(n, k, \ell)$  for  $n \approx k + \log_2 k$ . In particular, we have such graphs with  $n > k$ .

**Theorem 36** *Let  $\ell \geq 2$  and  $m \geq \max\{2\ell - 2, 4\}$ . A binary hypercube of dimension  $m$  belongs to  $\mathfrak{Gr}(2^m, 2^m - m + 2\ell - 2, \ell)$ .*

**Proof:** Suppose first  $\ell \geq 3$ . By [11, Theorem 2] we know that then a set in a binary hypercube is  $(1, \leq \ell)$ -identifying if and only if every vertex is dominated by at least  $2\ell - 1$  different vertices belonging to the set. Hence, we can remove any  $m + 1 - (2\ell - 1)$  vertices from the set of vertices, and there will still be a big enough multiple domination to assure that the remaining set is  $(1, \leq \ell)$ -identifying.

Suppose then that  $\ell = 2$  and  $G = (V, E)$  is the binary  $m$ -dimensional hypercube. Let us denote by  $C \subseteq V$  a  $(2^m - m + 2)$ -subset. Every vertex is dominated by at least  $m + 1 - (m - 2) = 3$  vertices of  $C$ . For all  $x, y \in V$ ,  $x \neq y$  we have  $|N[x] \cap N[y]| = 2$  if and only if  $1 \leq d(x, y) \leq 2$  and otherwise  $|N[x] \cap N[y]| = 0$ . Hence, for all  $x, y, z \in V$  with  $x \neq y$ ,  $I(y) = N[y] \cap C$  contains at least 3 vertices, and these cannot all be dominated by  $x$ ; thus, we have  $I(x) \neq I(y)$  and  $I(x) \neq I(y, z)$ .

We still need to show that  $I(x, y) \neq I(z, w)$  for all  $x, y, z, w \in V$ ,  $x \neq y$ ,  $z \neq w$ ,  $\{x, y\} \neq \{z, w\}$ . By symmetry we may assume that  $x \notin \{z, w\}$ . Suppose  $I(x, y) = I(z, w)$ .

If  $|I(x)| \geq 5$ , then any two vertices  $z, w \neq x$  cannot dominate  $I(x)$ , a contradiction.

If  $|I(x)| = 4$ , then  $|I(z) \cap I(x)| = |I(w) \cap I(x)| = 2$  and  $I(x) \cap I(z) \cap I(w) = \emptyset$ . It follows that  $3 \leq d(z, w) \leq 4$  which implies  $I(z) \cap I(w) = \emptyset$ . Since  $|N[x] \setminus C| = |N[x]| - |I(x)| = m - 3$ , all except one vertex, say  $v$ , of  $V \setminus C$  belong to  $N[x]$ , so  $V \setminus N[x] \subseteq C \cup \{v\}$ ; the vertex  $v$  cannot belong to both  $N[z]$  and  $N[w]$  since these are disjoint, so we may (w.l.o.g.) assume that  $v \notin N[z]$ , and thus  $N[z] \setminus N[x] \subseteq C$ , whence  $N[z] \setminus N[x] \subseteq I(z) \setminus I(x)$ . Hence,  $|I(z) \cap I(y)| \geq |I(z) \setminus I(x)| \geq |N[z] \setminus N[x]| = |N[z]| - |N[z] \cap N[x]| = m + 1 - 2 \geq 3$ . Thus  $y = z$ ; however, then  $I(y) \cap I(w) = I(z) \cap I(w) = \emptyset$  and since  $I(w) \not\subseteq I(x)$ , we have  $I(w) \not\subseteq I(x, y)$ .

Suppose finally that  $|I(x)| = 3$ ; w.l.o.g. we may assume  $|I(z) \cap I(x)| = 2$ . Now  $|N[x] \setminus C| = |N[x]| - |I(x)| = m - 2 = |V \setminus C|$ , and thus  $V \setminus C = N[x] \setminus C \subseteq N[x]$ ; hence,  $V \setminus N[x] \subseteq C$  and thus

$N[z] \setminus N[x] \subseteq I(z) \setminus I(x)$ . Consequently,  $|I(z) \cap I(y)| \geq |I(z) \setminus I(x)| \geq |N[z] \setminus N[x]| \geq m+1-2 \geq 3$ , and thus  $z = y$ . But similarly  $N[w] \setminus N[x] \subseteq I(w) \setminus I(x)$  and the same argument shows  $w = y$ , and thus  $w = z$ , a contradiction.  $\square$

We finally consider graphs without isolated vertices (i.e., no vertices with degree zero), and in particular connected graphs.

By [13, Theorem 8] a graph with no isolated vertices admitting a  $(1, \leq \ell)$ -identifying set has minimum degree at least  $\ell$ . Hence, always  $n \geq \ell + 1$ .

In [7] and [12] it has been proven that there exist connected graphs which admit  $(1, \leq \ell)$ -identifying sets. For example, the smallest known connected graph admitting a  $(1, \leq 3)$ -identifying set has 16 vertices [12]. It is unknown whether there are such graphs with smaller order. In the next theorem we solve the case of graphs admitting  $(1, \leq 2)$ -identifying sets.

**Theorem 37** *The smallest  $n \geq 3$  such that there exists a connected graph (or a graph without isolated vertices) in  $\mathfrak{Gr}(n, n, 2)$  is  $n = 7$ .*

(If we allow isolated vertices, we can trivially take the empty graph  $E_n$  for any  $n \geq 2$ .)

**Proof:** The cycle  $C_n \in \mathfrak{Gr}(n, n, 2)$  for  $n \geq 7$  by Example 3(ii) (see also [12]).

Assume that  $G = (V, E) \in \mathfrak{Gr}(n, n, 2)$  is a graph of order  $n \leq 6$  without isolated vertices; we will show that this leads to a contradiction. By [13], we know that  $\deg(v) \geq 2$  for all  $v \in V$ . We will use this fact frequently in the sequel.

If  $G$  is disconnected, the only possibility is that  $n = 6$  and that  $G$  consists of two disjoint triangles, but this graph is not even in  $\mathfrak{Gr}(n, n, 1)$ .

Hence,  $G$  is connected. Let  $x, y \in V$  be such that  $d(x, y) = \text{diam}(G)$ .

(i) Suppose that  $\text{diam}(G) = 1$ , or more generally that there exists a dominating vertex  $x$ . Then  $N[x, y] = N[x]$  for any  $y \in V$ , which is a contradiction.

(ii) Suppose next  $\text{diam}(G) = 2$ . Moreover, by the previous case we can assume that for any  $v \in V$  there is  $w \in V$  such that  $d(v, w) = 2$ .

Assume first  $|N(x)| = 4$ . Then  $S_2(x) = \{y\}$ . Since  $\deg(y) \geq 2$ , there exist two vertices  $w_1, w_2 \in N(y) \cap N(x)$ , but then  $N[x, w_1] = N[x, w_2]$ .

Assume next  $|N(x)| = 3$ , say  $N(x) = \{u_1, u_2, u_3\}$ . Then  $|S_2(x)| = n - |N[x]| \leq 2$ . Since the four sets  $N[x]$  and  $N[x, u_i]$ ,  $i = 1, 2, 3$ , must be distinct, we can assume without loss of generality that  $|S_2(x)| = 2$ , say  $S_2(x) = \{y, w\}$ , and that the only edges between the elements in  $S_2(x)$  and  $N(x)$  are  $u_1y, u_2w, u_3y$  and  $u_3w$ . Then  $N[x, u_3] = N[y, u_2]$ .

Assume finally that  $|N(x)| = 2$ . By the previous discussion we may assume that  $|N(v)| = 2$  for all  $v \in V$ . Then  $G$  must be a cycle  $C_n$ , but it can easily be seen that  $C_n \notin \mathfrak{Gr}(n, n, 2)$  for  $3 \leq n \leq 6$ .

(iii) Suppose that  $\text{diam}(G) = 3$ . Clearly  $|N(x)| \geq 2$  and  $|S_2(x)| \geq 1$ . If  $|S_2(x)| = 1$ , say  $S_2(x) = \{w\}$ , then  $N[w, y] = N[w]$ , which is not allowed. Since  $n \leq 6$ , we thus have  $|N(x)| = 2$  and  $|S_2(x)| = 2$ , say  $N(x) = \{u_1, u_2\}$  and  $S_2(x) = \{w_1, w_2\}$ . We can assume without loss of generality that  $u_1w_1 \in E$ . If  $w_2u_2 \in E$ , then  $N[w_1, u_2] = N[x, y]$ . If  $w_2u_2 \notin E$ , then  $N[w_1, w_2] = N[w_1]$ .

(iv) Suppose that  $\text{diam}(x, y) \geq 4$ . Then  $G$  contains an induced path  $P_5$ . There is at most one additional vertex, but it is impossible to add it to  $P_5$  and obtain  $\delta_G \geq 2$  and  $\text{diam}(G) \geq 4$ .

This completes the proof.  $\square$

## Acknowledgements

Part of this research was done during the Workshop on Codes and Discrete Probability in Grenoble, France. We would like to thank the referees for their comments.

**Tab. 1:** Lower and upper bounds for  $\Xi(k)$  for  $1 \leq k \leq 20$ . The lower bounds come from the examples given in the last column; for  $n \geq 8$  using Theorem 23, 25 or 27 or Lemma 10. The strongly regular graphs used here can be found from [5]. The upper bounds for  $k \geq 7$  come from Theorem 17.

k	lower bound	upper bound	example
1	1	1 (Ex. 12)	$E_1$
2	2	2 (Ex. 13)	$E_2$
3	4	4 (Ex. 14, Th.17)	$C_4, S_4$
4	5	5 (Th. 16)	Figure 1(b)
5	8	8 (Th. 17)	Example 9
6	9	9 (Th. 20)	Example 11, $P(9)$
7	11	12 (Th. 17, Th. 20)	Figure 1(a)
8	13	14	$P(13)$
9	16	16	RSHCD+
10	17	18	$P(17)$
11	18	20	Th. 25(ii)
12	21	22	(21,10,3,6)-SRG
13	22	24	Lemma 10
14	25	26	$P(25)$
15	26	28	(26,15,8,9)-SRG
16	29	30	$P(29)$
17	30	32	Th. 25(ii)
18	31	34	Th. 25(ii)
19	36	36	RSHCD+
20	37	38	$P(37)$



## References

- [1] D. Auger. Induced paths in twin-free graphs. *Electron. J. Combin.*, 15:N 17, 7 pp., 2008.
- [2] D. Auger, I. Charon, I. Honkala, O. Hudry, and A. Lobstein. Edge number, minimum degree, maximum independent set, radius and diameter in twin-free graphs. *Adv. Math. Commun.*, 3(1):97–114, 2009.
- [3] A. E. Brouwer, A. M. Cohen, and A. Neumaier. *Distance-regular graphs*, volume 18 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1989.
- [4] A. E. Brouwer and J. H. van Lint. Strongly regular graphs and partial geometries. In *Enumeration and design (Waterloo, Ont., 1982)*, pages 85–122. Academic Press, Toronto, ON, 1984.
- [5] C. J. Colbourn and J. H. Dinitz, editors. *The CRC handbook of combinatorial designs*. CRC Press Series on Discrete Mathematics and its Applications. CRC Press, Boca Raton, FL, 1996.
- [6] J.-M. Goethals and J. J. Seidel. Strongly regular graphs derived from combinatorial designs. *Canad. J. Math.*, 22:597–614, 1970.
- [7] S. Gravier and J. Moncel. Construction of codes identifying sets of vertices. *Electron. J. Combin.*, 12:R 13, 9 pp., 2005.
- [8] M. N. Huxley. *The Distribution of Prime Numbers*. Oxford University Press, London, 1972.
- [9] M. G. Karpovsky, K. Chakrabarty, and L. B. Levitin. On a new class of codes for identifying vertices in graphs. *IEEE Trans. Inform. Theory*, 44(2):599–611, 1998.
- [10] M. Laifenfeld and A. Trachtenberg. Identifying codes and covering problems. *IEEE Trans. Inform. Theory*, 54(9):3929–3950, 2008.
- [11] T. Laihonen. Sequences of optimal identifying codes. *IEEE Trans. Inform. Theory*, 48(3):774–776, 2002.
- [12] T. Laihonen. On cages admitting identifying codes. *European J. Combin.*, 29(3):737–741, 2008.
- [13] T. Laihonen and S. Ranto. Codes identifying sets of vertices. In *Applied algebra, algebraic algorithms and error-correcting codes (Melbourne, 2001)*, volume 2227 of *Lecture Notes in Computer Science*, pages 82–91. Springer, Berlin, 2001.
- [14] A. Lobstein. Identifying and locating-dominating codes in graphs, a bibliography. Published electronically at <http://perso.enst.fr/~lobstein/debutBIBidetlocdom.pdf>.
- [15] F. J. MacWilliams and N. J. A. Sloane. *The Theory of Error-Correcting Codes*, volume 16 of *North-Holland Mathematical Library*. North-Holland, Amsterdam, 1977.
- [16] W. Mantel. Problem 28. *Wiskundige Opgaven*, pages 60–61, 1907.
- [17] J. Seberry Wallis. Hadamard matrices. In W. D. Wallis, A. P. Street, and J. S. Wallis, editors, *Combinatorics: Room squares, sum-free sets, Hadamard matrices*, Lecture Notes in Mathematics, vol. 292, Springer-Verlag, Berlin-New York, 1972.